

---

**HMMER Sequences in the Rev Alignment**

---

A_U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
B_HXB2R	HIVHXB2R	K03455	Starcich,B.	Science 227, 484 (1985)
D_HIVELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
F_BZ163A	HIV1BZ163A	L22085	Louwagie,J.J.	ARHR 10, 561 (1994)
O_ANT70C	HIVANT70C	L20587	Vanden Haesevelde,M.	JVI 68, 1586 (1994)
O_MVP5180	HIVMVP5180	L20571	Gurtler,L.G.	JVI 68, 1581 (1994)
CPZGAB	SIVCPZGAB	X52154	Huet,T.	Nature 345, 356 (1990)
CPZANT	SIVCPZANT	U42720	Vanden Haesevelde,M.	Virology 221, 346 (1996)
A_ROD	HIV2ROD	M15390	Clavel,F.	Nature 324, 691 (1986)
B_EHOA	HIV2EHOA	U27200	Rey-Cuille,M.A.	Virology 202, 471 (1994)
SD_MM251	SIVMM251	M19499	Franchini,G.	Nature 328, 539 (1994)
STM_STM	SIVSTM	M83293	Novembre,F.J.	Virology 186, 783 (1992)
VER_AGM3	SIVAGM3	M30931	Baier,M.	Virology 176, 216 (1990)
GRI_AGM677	SIVAGM677	M66437	Fomsgaard,A.	Virology 182, 397 (1991)
SAB_SAB1C	SIVSAB1C	U04005	Jin,M.J.	EMBO J. 13, 2935 (1994)
SYK_SYK	SIVSYK	L06042	Hirsch,V.M.	JVI 67, 1517 (1993)

---

The following alignment and most-likely sequence were generated using the HMMER program as described in Part III. For simplicity, only representative types and subtypes are shown. The annotation is based on HIV1s, therefore the user should be cautious about its applicability to other PIV sequences. The high affinity binding site to RRE, which also contains a nuclear localization signal (nls) and the leucine effector domain are shown. The BLOCKMAKER program (using the Motif or Gibbs approach) produced one conserved, gapless block, which is shown by shading. The Motif "Cobbler" sequence follows this alignment.

	\/3'sj	exon\  exon	<- nls ->	
most-likely	MAGRSG..DSDEELLKTVRLIKLLYQ..	NPYPSPGTRQARRNRRRW.RERQRQIRSI	SARILSTFL	64
A_U455	--R---.NP--D---A--I-----	C-N-R-S----K-----	-A-----D-L-E----DC-	64
B_HXB2R	-----..IR-----	P-N-----	.H---E---G-Y-	64
D_ELI	-----..D---A---F-----	P-----	.A-----E-AE---G-Y-	64
F_BZ163A	-----..N-GT---RAA-Y--I-----	K-----	.A-----E--R---SC-	64
O_ANT70	-----..SE-DQ--QAIQI--I-----	Q--R-S-N--K-----	.R--A-VDTLA--V-A-VV	62
O_MVP5180	-----..EE-QQ--QAIQI--I-----	C-T-A-S-N--K-----	.R--A-VD-LAT--A-VV	63
CPZGAB	-----EP.QD-AR--QA-KI--I-----	K-----	.A--K--SE--G-V-A-Y-	65
CPZANT	-----EELEGT--Q---A-KI--I-----	K-A-S-A---K-----	KK--D-VEGLA--V-R-L	64
A_ROD	-NE-....ADE-G-QRKL--R--H-T.....	QGP--ASQ-----	KQ-W--L.AL-DSIY--.	60
B_EHOA	-NA-....ERD-Q-GL--LH--H-TSEYGT	QGP--ASQ-----	KQ-GL--L.AL-DRIHPL.	64
D_MM251	...S-H..ERE--R-RL--H--H-T.....	S--TGP--ANQ--Q-----	.R-WQ-LL.AL-DRIYS-.	60
STM_STM	...-...-QE--R-RL--QF-H-TT.....	D--QGP--ANQ-----	.Q-WN-LL.AL-NRIYS-.	59
VER_AGM3	...P...LGS--.RRLR--AF-NKN.....	PV--ARQ--RA-----	.QA-E-L-.AL-ERIWS.	57
GRI_AGM677	...-...LGK--.KQALKI--T--G-----	QFS--ARQ--RA-Q--.	KQ-Q--DK-AG-V-N--.	58
SAB_SAB1C	...-...LGQ---RRF-I--F--TT.....	PGQ--ARQ--RA-Q--AKQ--QV-H.LAE-----	56	
SYK_SYK	...P...-QGS-.QLA-F-RMIAHLQ.....	E---G---P--T--R---Q-.Q-RT-RL.YLQQRI....	54	
leucine-rich <-domain->				
most-likely	GRSAEPVP.LQLPPLERLTLDCE..	DCGTSGT.Q.....GVGSP.....QILVE.....		105
A_U455	--P-----I--R-----S-----	QPQGTET--G-----S-----		112
B_HXB2R	-----N-----			105
D_ELI	--P-----N-N-----R-----H-----	S-----		105
F_BZ163A	--PE-----P---INN---N-EQGAE.E-----		TSG-----	105
O_ANT70	HGPQNNNI.VD-----Q-SIRDP--.GDQL-EA.W.....T-DPRAEDNXCL-N-CS.....			110
O_MVP5180	HG-QDNNL.VD-----Q-NIRDP--.ADRLP--.G.....T-DPG.....TKDN			103
CPZGAB	--PPK-GD--E--E-DK-S-Q-V-TTQ-V--N-S-----	P...Q.....TATG-TVPA.....		111
A_ROD	...PD--.A..DS--DQT.IQH.L...QGL-IQE.L.....P.DP--.....TH-P-----			93
B_EHOA	...PDS--T..EG--DLA.IQR.L...QNLIIKD.L.....P.NP--.....TSTPTAQASTCIPPIWDQ			110
D_MM251	...PD--.T..DT--DLA.IQQ.L...QNLAIIES.I.....P.DP--.....TNTP-AL.....			95
STM_STM	...PD--.A..ST--DLA.VQQ.L...QGLSIQD.L.....P.DP--.....PN-PK.....			92
VER_AGM3	.-VE-Q.L.V-A..IDQ-V--Q....QHLAIQQ.L.....P.DP--.....SSS			89
GRI_AGM677	..EDQQQLVA--QE-QLENK-LVL...QHL.....P.DP--.....H-H			89
SAB_SAB1C	..ET--S.QID.H-AQEFDQLVL--.NL..QQ.P.....PSLP--.....GHPT--			91
SYK_SYK	...F-A.I.FGS..RTAA.-ED-L...QQLQISD			77
most-likely	.....SPTVLESG.....TK..E			116
A_U455	.....SA--G-----	N		123
B_HXB2R	.....-----			116
D_ELI	.....-----	EEQC		118
F_BZ163A	.....HA--G-----			116
O_ANT70	.....CN-I-ATR.....IA..-			121
CPZGAB	.....GGNYSI-GK-----	A..N		124
A_ROD	.....Q.R-----	AE..T		100
B_EHOA	LVPRSNPSSSQGCRDSCERGEDLVG--.Q--RRDHCNTQEDQ-R..G			156
D_MM251	.....CD--KGSRS.....PQ..D			107
STM_STM	.....D..QDT.....AE..N			100
SAB_SAB1C	.....NQ-ANS-S			99

## \*\*COBBLER sequence from MOTIF\*\*

>REV ELI, with embedded consensus blocks  
magrsgdsdedLLRAIRLIKLLYQSNPYQPEGTRQQRRNRRWRQRQRQIxALAERIL  
STylgrpaevpvlqlpplerlnlncsedcrtsgtqvgvghpqisvesptvlesgtteeqc

## HIV1 REV CONSENSUS

	high-affinity binding site	
	nls	
	\ / 3' sj                    exon \ / exon     / <-                    -> /	
CONSENSUS-A	MAgRSG? sDE . eLL . KAI RII KIL YQSNP yPK PkG . SRQARKNRR RWR QRID S1Se RIL St CLGRP	66
CONSENSUS-B	-----d---.---. -tV-1--f-----p-s-e-.T---R-----e----r-i--w----y---s	67
CONSENSUS-F	-----N-?T.---.R-?-Y-----E-.T---R-----? -R??-?----S-----	61
CONSENSUS-O	-----E-...Q?-?Q--Q-----? -? -? .--N-----R--A-V-? -A?-? -A-VH G?	56
CONSENSUS-U	-----DA---.RVV-----P-E-.T--T-----RAI---F-----S	67
CONSENSUS-CPZ	---?E-?????? -? ? -VK-----? ---? -? . ? -? --R-? ---? -? -? ? -? ---V-? -? ---	41
	Leu-rich effector domain	
	/ <-                    -> /	
CONSENSUS-A	AEPVPLQLPP1ERLhLDCsEdcgTSgTQq?qg?etGVGrpQvsVEssavLGSGTkn	120
CONSENSUS-B	-----t---?-----.....?---s--il---p---e----E\$	115
CONSENSUS-F	E-----?---?IN?--? -E.Q-A?E.....---S--T-G--H-----E\$	105
CONSENSUS-O	Q?NN?VD-----Q-?IRDP-?D?L????TVDPRAEDN\$CL-NLCSCNT???????N\$	95
CONSENSUS-U	-----I---C-----G-----P-T-----S-PI-G---TI-----E\$	123
CONSENSUS-CPZ	PK-GD-E--E-DK-S-Q-V-TTQDV--SNTSQPQ-AT-ETVPAGGNYSI--K-A--	97